

IMAGE RETRIEVAL DEVICE, IMAGE RETRIEVAL METHOD AND STORAGE MEDIUM
STORING SIMILAR-IMAGE RETRIEVAL PROGRAM

BACKGROUND OF THE INVENTION

5

Field of the Invention

The present invention relates to an image retrieval device
and an image retrieval method to retrieve an image being similar
to an inquired image, from images accumulated in an image database
10 and a storage medium storing an image retrieval program to
retrieve, by controlling a computer, an image being similar to
an inquired image from images stored in an image database

The present application claims priority of Japanese Patent
15 Application No. Hei 11-351657 filed on December 10, 1999, which
is hereby incorporated by reference.

Description of the Related Art

20 There are cases where an image being similar to a specified
image (hereinafter referred to as an "inquired image") is
retrieved from an image database in which image data is
accumulated and managed. Many conventional similar-image
retrieving technologies of this kind have been proposed. In such
25 conventional technologies, color information is mainly used as
an image feature descriptor. Moreover, in most of the conventional
technologies, a histogram of color information and a comparison
process based on similarity obtained by the histogram is employed
to retrieve a targeted image. However, this method has a problem

09732701-121100

in that a color structure of an image is not reflected.

To solve this problem, similar-image retrieval technologies are disclosed in, for example, Japanese Patent Application Laid-open Hei 8-249349 in which the image is retrieved with high accuracy by introducing an attempt to have the color structure of the image reflected in the image retrieval and where, in an image database, the image is divided into a plurality of blocks and then a typical color of each block is calculated as an image feature descriptor and a pattern matching is performed.

However, in this conventional technology, since the typical color of each block is calculated, a scale of the image feature descriptor is made larger, causing a reduction in a retrieval speed. This also causes a size of a hardware required for retrieval processing to be made larger.

Another conventional method to solve this problem is to perform an orthogonal transform on the image to efficiently express the image feature. Figure 4 is a schematic block diagram showing configurations of main parts of a conventional image retrieval device using a transform coefficient. As shown in Fig. 4, an image feature descriptor storing section 41 stores a transform coefficient 401 for various image data. A similarity calculating section 42 compares a transform coefficient 403 contained in inquired image data with an image feature descriptor 402 accumulated in the image feature descriptor storing section 41 to calculate similarity between them and outputs its result 404.

By performing an orthogonal transform of an image and by using a part of the coefficient as an image feature descriptor,

a scale of the image feature descriptor can be made smaller. This allows the retrieval processing to be made high-speed and a size of a hardware to be made smaller.

Furthermore, another method to improve image retrieval accuracy is to decode a transform coefficient. Figure 5 is a schematic block diagram showing configurations of another conventional image retrieval device which performs decoding of the transform coefficient. As shown in Fig. 5, the conventional image retrieval device is composed of an image feature descriptor storing section 51, an inverse orthogonal transforming device 52, a color space transforming device 53 and a similarity calculating section 54. The image feature descriptor storing section 51 stores, in advance, a transform coefficient 501 of image data as an image feature descriptor. The inverse orthogonal transforming device 52 performs an inverse orthogonal transform on an image descriptor 502 accumulated in the image feature descriptor storing section 51 and outputs transformed image data 503. The color space transforming device 53 transforms color space of image data 504 output from the inverse orthogonal transforming device 52. Inquired image data 505 is input as image data 506 whose color space is transformed by a color space transforming device 55 to the similarity calculating section 54. The similarity calculating section 54 calculates similarity between the input image data 506 and the image data 504 obtained from the color space transforming device 53 and outputs its result 507.

However, each of above conventional similarity image retrieving technologies has following problems:

In the conventional technology to try to express

09732701.121100

efficiently an image feature by performing an orthogonal transform on an image, though retrieval processing can be made high-speed and a size of a hardware required for image retrieval can be made smaller, the conventional technology cannot detect
5 an image having a complete visual similarity or detects an image which is not visually similar in some cases.

This is because there is no conformity in a distance between images expressed by transform coefficients and visual similarity among images, causing insufficient retrieval accuracy.

10 In the conventional technology to match an image obtained by reconstructing the image by decoding orthogonally-transformed coefficient and then mapping the image over color space such as a HSV (Hue, Saturation, Value) or a like, with an inquired image, though excellent retrieval accuracy can be obtained, the image
15 retrieval is very costly and speed of the image retrieval is reduced. This is because it is necessary to perform decoding processing and color space transforming processing, at every time of the retrieval, on each of image feature descriptors of accumulated data.

20

SUMMARY OF THE INVENTION

In view of the above, it is an object of the present invention to provide an image retrieval device, an image retrieval method
25 and a storage medium storing similar-image retrieval program which are capable of simplify system configurations without incurring failure of image retrieval performance.

Also, it is another object of the present invention to

0922701-121100

provide an image retrieval device, an image retrieval method and a storage medium storing similar-image retrieval program which are capable of retrieving images at higher speed with simplified calculating processes.

5 According to a first aspect of the present invention, there is provided an image retrieval device for retrieving an image being similar to an inquired image from images stored in an image database including:

10 a first coefficient transforming means for transforming a first group of image feature descriptors extracted from image data accumulated in the image database and then generating a second group of image feature descriptors to be used for calculating similarity;

15 a second coefficient transforming means for transforming a first group of image feature descriptors extracted from image data of the inquired image and then generating a second group of image feature descriptors to be used for calculating similarity; and

20 a similarity calculating means for calculating similarity by comparing the second group of image feature descriptors for each piece of image data generated by the first coefficient transforming means with the second group of image feature descriptors transformed by the second coefficient transforming means.

25 In the foregoing, a preferable mode is one that wherein includes an image feature descriptor storing means and wherein the similarity calculating means compares the second group of image feature descriptors of image data of the inquired image

00732701-121100

received from the second coefficient transforming means with the second group of image feature descriptors of image data contained in the image database read from the image feature descriptor storing means.

5 Also, a preferable mode is one wherein the first coefficient transforming means and the second coefficient transforming means perform transform of image feature descriptors in a manner that visual similarity between images to be compared is approximated by a distance between an image expressed by the second group of
10 image feature descriptors of image data contained in the image database and an image expressed by the group of image feature descriptors of image data of the inquired image.

Also, a preferable mode is one wherein the first coefficient transforming means and the second coefficient transforming means
15 use, as the image feature descriptor, a transform coefficient obtained by performing specified transforming processing of coefficient on image data.

Also, a preferable mode is one wherein the first coefficient transforming means and the second coefficient transforming means
20 perform transform of the coefficient using a transform table selected depending on a kind of transform coefficient to be used as the image feature descriptor and wherein the first coefficient transforming means and the second coefficient transforming means perform retrieval of a similar-image on a trial basis using a
25 plurality of the transform tables each having a differently segmented range of the transform coefficient and select a transform table which has showed a high rate of correctly solved retrieval in the retrieval on a trial basis.

09732701-121100

Also, a preferable mode is one wherein the first coefficient transforming means and the second coefficient transforming means perform a transform of the coefficient in a manner that fine quantization is carried out on a range of a portion of each of the transform coefficients having a small amplitude and that coarse quantization is carried out on a range of a portion of each of the transform coefficients having a large amplitude.

Furthermore, a preferable mode is one wherein the first coefficient transforming means and the second coefficient transforming means perform transform of the coefficient in a manner that fine quantization is carried out on a coefficient having a small power and coarse quantization is carried out on a coefficient having a large power.

According to a second aspect of the present invention, there is provided an image retrieval device for retrieving an image being similar to an inquired image from images stored in an image database including:

a coefficient transforming means for transforming a first group of image feature descriptors extracted from image data accumulated in the image database and from image data of the inquired image and then generating a second group of image feature descriptors used to calculate similarity; and

a similarity calculating means for comparing the second groups of image feature descriptors generated by the coefficient transforming means and then calculating similarity between an image accumulated in the image database and the inquired image.

According to a third aspect of the present invention, there is provided an image retrieving method for retrieving an image

09732701.121100

being similar to an inquired image from images stored in an image database, the method including:

5 a step of transforming a first group of image feature descriptors extracted from image data accumulated in the image database and then generating a second group of image feature descriptors to be used for calculating similarity;

10 a step of transforming a first group of image feature descriptors extracted from image data of the inquired image and then generating a second group of image feature descriptors to be used for calculating similarity; and

a step of comparing the second group of image feature descriptors of image data accumulated in the image database with the second group of image feature descriptors of image data of the inquired image to calculate similarity.

15 In the foregoing, it is preferable that, in the two steps in which the first group of image feature descriptors is transformed and the second group of image feature descriptors is generated, a transform of the coefficient is performed using a transform table selected depending on a kind of transform
20 coefficient to be used as the image feature descriptor and wherein retrieval of a similar-image is performed on a trial basis using a plurality of the transform tables each having a differently segmented range of the transform coefficient and a transform table is selected which has showed a high rate of correctly solved
25 retrieval in the retrieval on a trial basis.

According to a fourth aspect of the present invention, there is provided a storage medium storing a similar-image retrieval program to cause a computer to carry out retrieval of an image

09732701-121100

being similar to an inquired image from images stored in an image database, wherein the similar-image retrieval program includes a step of transforming a first group of image feature descriptors extracted from image data accumulated in the image database and
5 then generating a second group of image feature descriptors to be used for calculating similarity, a step of transforming a first group of image feature descriptors extracted from image data of the inquired image and then generating a second group of image feature descriptors to be used for calculating similarity; and
10 a step of comparing the second group of image feature descriptors of image data accumulated in the image database with the second group of image feature descriptors of image data of the inquired image to calculate similarity.

In the foregoing, a preferable mode is one that wherein
15 includes an image size transforming means for increasing or decreasing image data and/or inquired image data accumulated in the one wherein the image feature descriptor producing means performs a discrete cosine transform (DCT) on an image obtained by the image size transforming means and extracts an obtained DCT
20 coefficient and uses the DCT coefficient as a first group of image feature descriptors.

Also, a preferable mode is one that wherein includes an image size transforming process of increasing or decreasing image data and/or inquired image data accumulated in the image database in
25 size and an image feature descriptor producing process of performing an orthogonal transform on an image obtained by the image size transforming process and producing an orthogonal transform coefficient and using the orthogonal transform

09732701-121100

coefficient as a first group of image feature descriptor image database in size and an image feature descriptor producing means for performing an orthogonal transform on an image obtained by the image size transforming means and producing an orthogonal transform coefficient and using the orthogonal transform coefficient as a first group of image feature descriptors.

Also, a preferable mode is one wherein the image size transforming means has a block dividing means for dividing the image data into blocks, a typical color calculating means for calculating a typical color of each of blocks obtained by the block dividing means and an image creating means for creating an image using the typical color of each of the blocks as a pixel.

Also, a preferable mode is one wherein the image creating means extracts a color average of entire pixels contained in each of the blocks as the calculated typical color of each of the blocks.

Also, a preferable mode is one wherein the block dividing means divides the image data into 64 blocks.

Also, a preferable mode is one wherein the image size transforming process has a block dividing process of dividing the image data into blocks, a typical color calculating process of calculating a typical color of each of blocks obtained by the block dividing process and an image creating process of creating an image using the typical color of each of the blocks as a pixel.

Also, a preferable mode is one wherein the image creating process is to extract a color average of entire pixels contained in each of the blocks as the calculated typical color of each of the blocks.

Also, a preferable mode is one wherein the block dividing

09732701-124100

process is to divide the image data into 64 blocks.

Furthermore, a preferable mode is one wherein the image feature descriptor producing process is to perform a DCT on an image obtained by the image size transforming process and to extract an obtained DCT coefficient and to use the DCT coefficient as a first group of image feature descriptors.

With the above configurations, since second image feature descriptors to be used for calculating similarity are computed by transforming first image feature descriptors extracted, in advance, from image data and then similarity is calculated by directly using the second image feature descriptors, decoding processing on image feature descriptors and pattern matching at a time of retrieval of a similar-image are not required. This allows retrieval processing to be performed with simplified calculating processes and configurations of the retrieval device to be made simple and compact.

Moreover, since the first image feature descriptor is transformed to the second image feature descriptor in a manner that visual similarity between images is approximated by a distance between images expressed by groups of the second image feature descriptors, retrieval performance can be successfully improved.

Therefore, it is possible that retrieving images successfully at higher speed with simplified configurations.

BRIEF DESCRIPTION OF THE DRAWINGS

The above and other objects, advantages and features of the

present invention will be more apparent from the following description taken in conjunction with the accompanying drawings in which:

Fig. 1 is a schematic block diagram showing configurations of an image retrieval device according to a first embodiment of the present invention;

Fig. 2 is a flowchart explaining selecting procedures in a transform table used in the image retrieval device according to the first embodiment of the present invention;

Fig. 3 is a table showing a comparison of a rate of correctly solved retrieval for same database and inquired image obtained by using a DCT (Discrete Cosine Transform) coefficient itself, a transformed DCT coefficient or image decoding according to the first embodiment of the present invention;

Fig. 4 is a schematic block diagram for showing configurations of main parts of a conventional image retrieval device using a transform coefficient; and

Fig. 5 is a schematic block diagram for showing configurations of another conventional image retrieval device.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Best modes of carrying out the present invention will be described in further detail using various embodiments with reference to the accompanying drawings.

First Embodiment

09732701-121100

5
10

15
20
25

optical disks, semiconductor memories and other general storage media.

The coefficient transforming sections 11 and 13 converts each factor of a first group of image feature descriptors
5 extracted from image data to each factor of a second group of image feature descriptors to be used for calculation of similarity to be performed by the similarity calculating section 14. In this embodiment, a transform coefficient is used as the image feature descriptor. The transform coefficient represents a coefficient
10 calculated by performing a Discrete Cosine Transform (DCT), wavelet transform, Hadamard Transform or a like on image data. A transform can be achieved by applying an individual transform method to each of coefficients of image feature descriptors or by applying a common transform method to two or more coefficients.

15 Moreover, as described above, the coefficient transforming section 11 extracts a first image feature descriptor 101 from image data accumulated in the image database and converts and then stores a second image feature descriptor 102 obtained by the transformation in the image feature descriptor storing section
20 12. The coefficient transforming section 13 extracts a first image feature descriptor 103 from image data of inquired image and converts and feeds a second image feature descriptor 104 obtained by the transformation to the similarity calculating section 14.

The coefficient transforming sections 11 and 13 perform a
25 transform of the coefficient in a manner that visual similarity among images is approximated by a distance between images expressed by each of the second image feature descriptors 102, 104. That is, the transform of the coefficient is performed so

09732701-124100

that the distance between images expressed by each of the second image feature descriptors 102, 104 in both images is made smaller in an image having visual similarity and the distance between images expressed by each of the second image feature descriptors 102, 104 in both images is made larger in an image having no visual similarity. There is no limitation to a type of the transform method so long as such a result as above is obtained. One example of it is a method using a transform table. The method using the transform table will be described in detail later.

Moreover, in the first embodiment, though the coefficient transforming section 11 to transform the first image feature descriptor 101 stored in an image database to the second image feature descriptor 102 and the coefficient transforming section 13 to transform the first image feature descriptor 103 extracted from image data of an inquired image are separately disposed, in an actual device, these two coefficient transforming sections 11 and 13 may be implemented as a same processing unit which may be configured in a manner that the unit judges whether input data is image data contained in an image database or the input data is image data contained in an inquired image and decides a place to which a produced second image feature descriptor 102, 104 is transferred to.

The image feature descriptor storing section 12 stores the second image feature descriptor 102 of image data stored in an image database which has been transformed by the coefficient transforming section 11. The image feature descriptor storing section 12, in response to a request from the similarity calculating section 14, sends a second image feature descriptor

09732701.121100

105 to the similarity calculating section 14.

5 The similarity calculating section 14 compares a second image feature descriptors 104 of an inquired image received from the coefficient transforming section 13 with the second image feature descriptor 105 received from the image feature descriptor storing section 12 to calculate similarity between them. That is, a distance between images expressed by each of the two second image feature descriptors 104, 105 is calculated. An image having high similarity, that is, the image having a distance being near to
10 an image expressed in the second image feature descriptor 104 obtained from the inquired image, out of images expressed by the second image feature descriptor 105, is selected and the image contained in an image database corresponding to the second image feature descriptor 105 is detected as an image 106 being similar
15 to the inquired image. Moreover, similarity and non-similarity between images can be determined by setting an appropriate threshold and by judging whether a distance between images expressed by the two second image feature descriptors 104 and 105 is larger or smaller than the threshold.

20 The similarity calculating section 14, as described above, calculates similarity using the second image feature descriptor 104 obtained by the inquired image and the second image feature descriptor 104 stored in the image feature descriptor storing section 12. Therefore, when a similar-image is retrieved,
25 decoding processing of the second image feature descriptor 104 is not required. This enables configurations of the retrieval device to be simplified. Moreover, since the second image feature descriptor 104 stored in advance is directly used to calculate

093301-12100

similarity, retrieval processing can be performed by simple calculating processes.

Next, the method in which a transform table used for a process of transforming an image feature descriptor in the coefficient transforming sections 11 and 13 is selected will be described below. Figure 2 is a flowchart explaining a selecting procedure in the transform table used in the image retrieval device according to the first embodiment. As shown in Fig. 2, image data and a transform coefficient corresponding to the image data is first prepared for initial setting. A transform table corresponding to the transform coefficient is prepared (Step 201).

As described above, various transform methods to provide transform coefficients may be used, which allows a variety of transform tables to be selected. It is necessary to prepare a transform table in which a coefficient is transformed in a manner that a range of a portion of each of transform coefficients having a small amplitude is finely segmented and a value is assigned to each of the segmented ranges, that is, the range of each of the transform coefficients is quantized and in a manner that a range of a portion of each of transform coefficients having a large amplitude is coarsely segmented and a value is assigned to each of the segmented ranges, that is, a range of each of the transform coefficients is quantized. Furthermore, it is necessary to prepare a transform table in which a coefficient is transformed in a manner that a range of each of two or more coefficients having a small power is finely segmented and a value is assigned to each of the segmented ranges, that is, the range of each of the two

09732701-121100

or more coefficients is finely quantized and in a manner that a range of each of two or more coefficients having a large power is coarsely segmented and a value is assigned to each of the segmented ranges. Therefore, a plurality of tables each having
5 a slightly different range to be segmented is prepared. Either of a transform table in which a coefficient is transformed based on an individual standard for each coefficient or a transform table in which two or more coefficients are transformed based on a same standard can be used.

10 Next, an inquired image and an image being visually similar to the inquired image which has been accumulated in image database are set (Step 202). In application, an image being visually similar to the inquired image is called a "correctly solved image". In order to improve retrieval accuracy, a plurality of inquired
15 images and correctly solved images which can correspond to each of the inquired images are set and following processing is performed on each of the correctly solved images. Moreover, a judgement as to which image is visually similar to the inquired image depends on a subjective judgement of an operator.

20 Next, the coefficient transforming sections 11 and 13 read the prepared transform table and calculate a transform coefficient of image data and then perform a coefficient transform using the read transform table (Step 203). The coefficient transforming section 11 calculates a transform coefficient for
25 each of image data accumulated in image database and performs a transform using the calculated transform coefficient. The coefficient transforming section 13 calculates a transform coefficient of image data of the inquired image to perform a

09732701-121100

transform. The coefficient transformed by the coefficient transforming section 11 is stored in the image feature descriptor storing section 12. The similarity calculating section 14 compares a coefficient (second image feature descriptor) obtained
5 by transforming an inquired image with a coefficient (second image feature descriptor) obtained by transforming stored in the image feature descriptor storing section 12 to calculate similarity (Step 204). Then, each of images stored in the image database is sorted in decreasing order of similarity (Step 205) and an average
10 value of orders in which a correctly solved image is detected (Step 206).

An average value of orders in which a correctly solved image is detected by performing procedures Step 204 to Step 206 for each of all inquired images is calculated (Step 207) and then an average
15 value of calculated average orders (hereinafter referred to as "retrieval average orders") in which a correctly solved image for each of all inquired images is calculated (Step 208).

Next, the retrieval average orders calculated at Step 208 and transform tables are registered on the coefficient
20 transforming sections 11 and 13. When other transform tables and retrieval average orders have been already registered, a retrieval average order newly obtained is compared with the retrieval average orders already registered. If the retrieval average order newly obtained is larger than the retrieval average
25 order already registered, the registered transform table and retrieval average order are maintained as they are. If the retrieval average order newly obtained is smaller than the retrieval average order already registered, the registered

00732701-124100

transform table and retrieval average order are replaced with a new transform table and a retrieval average order (Steps 209 and 210).

5 Same processing as above is sequentially performed on the two or more prepared transform tables and a retrieval average order is obtained in the transform table. The transform table having a small retrieval average order and retrieval average order remain left. When the calculation of the retrieval average order in all the prepared tables is terminated, a stored table, that is, a transform table from which a minimum retrieval average order is obtained is output as an optimum transform table and the processing is terminated (Step 211).

15 Moreover, though, in the above embodiment, an average value of orders in which a correctly solved image is retrieved is used as an evaluation value, a transform table in which an average value of similarity obtained when a correctly solved image is retrieved is used as an evaluation value and in which a sum of the average value of similarity to an inquired image is maximized, may be used.

20 Next, results from actual retrieval of a similar-image are shown below. First, by carrying out an experiment using an image database including 2,045 pieces of images, a coefficient transform table is created. Then, images (correctly solved images) evaluated, based on a subjective evaluation of the operator, to be similar to 26 pieces of inquired images are selected in advance. A value obtained by adding, with weights assigned, squared errors of each of transformed coefficients is employed as a distance between images and the distance is sorted in increasing order of distance and an average of orders in which

00772701-12180

a correctly solved image is retrieved (retrieval average order) is used as an evaluation value. Then, an evaluation is performed by using a plurality of transformed coefficient tables and weighted coefficients and a transform coefficient table in which
5 the retrieval average order is lowest is calculated.

Image feature descriptors were generated by the following methods. First, a still image is divided into 8 x 8 blocks and an average color of each block is calculated and a resized image having a fixed size (8 pixels x 8 pixels) is generated. Then, a
10 DCT transform is performed on the resized image and a set of low-degree coefficients expressing luminance and a chrominance signal, out of obtained coefficients, is extracted. Furthermore, the obtained DCT coefficients are transformed by using a transform table and the transformed coefficients are used as image feature
15 descriptors.

By using an image database containing 5,466 pieces of images and a most suitable transform table calculated on the above conditions, an experiment of retrieval of similar-images was carried out.

20 Images each being evaluated, from a subjective viewpoint, to be similar to each of 50 types of inquired images, are selected in advance and the selected images are defined to be correctly solved images. Sums of weighted squared difference between image feature descriptors of images contained in inquired images and
25 in database are sorted in decreasing order of difference and the number of correctly solved images ranked within the n-th position of high orders are evaluated to be a rate of correctly solved retrieval. The rate of correctly solved retrieval is defined to

09732701-121100

be a quotient obtained by dividing all the number of correctly solved images by the number of correctly solved images ranked within the n-th position of high orders and the "n" is defined to be a value being four times larger than the number of correctly
5 solved images selected in advance.

Figure 3 is a table showing a comparison of the rate of correctly solved retrieval for same database and inquired image obtained by using the conventional technology in which a similar-image is retrieved by a DCT coefficient itself, by using
10 this embodiment of the present invention in which a similar-image is retrieved by a transformed DCT coefficient or by using the conventional technology in which a similar-image is retrieved by decoding after the DCT transform and performing pattern matching. As is apparent from Fig. 3, by using the present embodiment in
15 which the transform table is introduced, the rate of correctly solved retrieval is greatly improved. An almost same performance as obtained by decoding an 8 x 8 image and performing pattern matching can be achieved by using a far less number of coefficients.

20 It is apparent that the present invention is not limited to the above embodiments but may be changed and modified without departing from the scope and spirit of the invention.

09732701-121100